UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/585,235	06/29/2006	Guodong Zhou	284502000900	9355

20872 7590 02/23/2010

MORRISON & FOERSTER LLP
425 MARKET STREET
SAN FRANCISCO, CA 94105-2482

EXAMINER	
COLUCCI, MICHAEL C	

ART UNIT	PAPER NUMBER
2626	

MAIL DATE	DELIVERY MODE
02/23/2010	PAPER

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Office Action Summary	Application No. 10/585,235	Applicant(s) ZHOU ET AL.	
	Examiner MICHAEL C. COLUCCI	Art Unit 2626	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☐ Responsive to communication(s) filed on _____.

2a)☐ This action is **FINAL**. 2b)☒ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *1-23* is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *1-23* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☒ The drawing(s) filed on *29 June 2006* is/are: a)☒ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☒ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☒ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☒ Information Disclosure Statement(s) (PTO/SB/08) Paper No(s)/Mail Date _____.
4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____ .
5) ☐ Notice of Informal Patent Application
6) ☐ Other: _____.

DETAILED ACTION

Claim Rejections - 35 USC § 101

1. 35 U.S.C. 101 reads as follows:

> Whoever invents or discovers any new and useful process, machine, manufacture, or composition of matter, or any new and useful improvement thereof, may obtain a patent therefor, subject to the conditions and requirements of this title.

Claims 1-23 are rejected under 35 U.S.C. 101 because:

Claims 1-23 do not fall within one of the four statutory categories of invention. Supreme Court precedent[1] and recent Federal Circuit decisions[2] indicate that a statutory "process" under 35 U.S.C. 101 must (1) be tied to another statutory category (such as a particular apparatus), or (2) transform underlying subject matter (such as an article or material) to a different state or thing. While the instant claim(s) recite a series of steps or acts to be performed, the claim(s) neither transform underlying subject matter nor positively tie to another statutory category that accomplishes the claimed method steps, and therefore do not qualify as a statutory process.

Claims 1-23 recite purely mental steps and would not qualify as a statutory process. In order to qualify as a statutory process, the method claim should positively recite the other statutory class to which it is tied (i.e. apparatus, device, product, etc.). For example, the system steps of claim 18 appear to recite mental steps such as "a system for recognizing and classifying named entities within a text". For example a

---

[1] *Diamond v. Diehr*, 450 U.S. 175, 184 (1981); *Parker v. Flook*, 437 U.S. 584, 588 n.9 (1978); *Gottschalk v. Benson*, 409 U.S. 63, 70 (1972); *Cochrane v. Deener*, 94 U.S. 780, 787-88 (1876).

[2] *In re Bilski*, 88 USPQ2d 1385 (Fed. Cir. 2008).

human can use a computer to read a portion of text and mentally perform feature extraction, classification, modeling, and even constraint relaxation. For instance, there is no indication in the claims that feature extraction, classification, modeling, and constraint relaxation *are performed in the computer system.* These steps do not identify an apparatus that performs the recited method steps, such as but not limited to, the system with hardware and software modules as described in the specification (present invention spec. page 26).

Claim Rejections - 35 USC § 103

2. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

3. Claims 1-23 are rejected under 35 U.S.C. 103(a) as being unpatentable over Zhou et al., "Named Entity Recognition using an HMM-based Chunk Tagger", July 2002 (hereinafter Zhou) in view of Taira US 20030105638 A1 (hereinafter Taira).

Re claims 1 and 12, Zhou teaches a method of back-off modeling for use in named entity recognition of a text, comprising, for an initial pattern entry from the text (Abstract):

valid and invalid forms of constraints (page 475 col. 2)

However, Zhou fails to teach relaxing one or more constraints of the initial pattern entry; determining if the pattern entry after constraint relaxation has a valid form; and

moving iteratively up the semantic hierarchy of the constraint if the pattern entry after constraint relaxation is determined not to have a valid form

Taira initially teaches matching text segments where statistics are accumulated on the frequency of occurrence within the training set of how often a word class is associated with a given logical relation role as head, relation, or value. This model is created to allow for the rapid indexing of the possible logical relations in which each word might be involved. Thus, information is gathered about the probabilistic context-free role of the particle in constructing logical relations. This initial characterization is used as a first pass filter to determine possible candidate logical relations for a given word (Taira [0107]).

Further, Taira teaches a degree of deformability is desirable in a template matching algorithm whose goal is high recall. For example, some semantic smoothing can be performed so that individual slot value constraints of the logical relation template are relaxed (Taira [0108]).

Furthermore, Taira teaches the relaxing of constraints to achieve a match if frequency of occurrence does not produce an exact match, wherein Taira teaches [0111] One problem inherent in creating a template matching algorithm automatically from a finite set of training data is that it is unlikely that every combination of words that might be used in a natural language medical report will be addressed. Unforeseen

combinations of words will leave gaps in the systems ability to identify all possible logical relation word pairs consisting of heads and values or triplets consisting of heads, relations, and values. To account for this problem, if there is no logical relation slot assignment for a particular words, one will be estimated using a graded relaxation of the syntactic and semantic properties of the word. FIG. 12 illustrates the mechanism used for relaxation. The semantic classes assigned to words is organized into a type abstraction hierarchy. This hierarchy allows a word such as "mass" to be relaxed to the more general abstract class "physobj.abnormal.findin- g.lesion". Within this class, the terms "nodule", and "cyst" are included. In this way, the restrictions of what types of words can fill specific roles of a given logical relation are relaxed. Relaxation can occur as many times as necessary in order to commit each word within the sentence to at least one possible logical relation role. Because there are hundreds of thousands of words in the English language, therefore it is likely that some words will not have appeared together in sentences in the training corpora and, as a result, there will be no template slot assignments for them. On the other hand, because the number of unique semantic classes is on the order of 500 and the number of syntactic classes is on the order of 15, these numbers are both finite sufficiently small to ensure that at least a default role can be supplied for any unforeseen combination of words. Moreover, of the 500 semantic classes, many of these are parts of hierarchies. For example, "mass" is classified as "physobj.abnorm.finding.lesion," a class which can be broadened to "physobj.abnorm.finding" or broadened even further to the class of all "physobj.abnorm." Accordingly, the builders of the database can choose how they might want to balance

accuracy and detail in building this concept database (Taira [0111-0112] & Fig. 7 iterations, Fig. 12 relaxing and broadening terms).

Therefore, it would have been obvious to one of ordinary skill in the art at the time of the invention to modify the system of Zhou to incorporate relaxing one or more constraints of the initial pattern entry; determining if the pattern entry after constraint relaxation has a valid form and moving iteratively up the semantic hierarchy of the constraint if the pattern entry after constraint relaxation is determined not to have a valid form as taught by Taira to allow for identification of the best match of input NER data with various interpretations, wherein relaxation can occur as many times as necessary in order to commit each word within the sentence to at least one possible logical relation role (Taira [0111-0112] & Fig. 7 iterations, Fig. 12 relaxing and broadening terms)., thereby valid constraint forms as taught by Zhou can be improved to include a broad and narrow use of NER entries by continuously broadening (if necessary) through disambiguation, frequency of occurrence, and constraint relaxation (Taira Fig. 7 for example iterative elements 730, 740, and 750, having relaxation applied [0111-0112]).

Re claims 2 and 13, Zhou teaches valid and invalid constraint forms in a hierarchy of chunks (page 475 col. 2).

However, Zhou fails to teach moving iteratively up the semantic hierarchy of the constraint if the pattern entry after constraint relaxation is determined not to have a valid form comprises moving up the semantic hierarchy of the constraint, relaxing the

constraint further, and returning to determining if the pattern entry after constraint relaxation has a valid form

Taira initially teaches matching text segments where statistics are accumulated on the frequency of occurrence within the training set of how often a word class is associated with a given logical relation role as head, relation, or value. This model is created to allow for the rapid indexing of the possible logical relations in which each word might be involved. Thus, information is gathered about the probabilistic context-free role of the particle in constructing logical relations. This initial characterization is used as a first pass filter to determine possible candidate logical relations for a given word (Taira [0107]).

Further, Taira teaches a degree of deformability is desirable in a template matching algorithm whose goal is high recall. For example, some semantic smoothing can be performed so that individual slot value constraints of the logical relation template are relaxed (Taira [0108]).

Furthermore, Taira teaches the relaxing of constraints to achieve a match if frequency of occurrence does not produce an exact match, wherein Taira teaches [0111] One problem inherent in creating a template matching algorithm automatically from a finite set of training data is that it is unlikely that every combination of words that might be used in a natural language medical report will be addressed. Unforeseen combinations of words will leave gaps in the systems ability to identify all possible logical relation word pairs consisting of heads and values or triplets consisting of heads,

relations, and values. To account for this problem, if there is no logical relation slot assignment for a particular words, one will be estimated using a graded relaxation of the syntactic and semantic properties of the word. FIG. 12 illustrates the mechanism used for relaxation. The semantic classes assigned to words is organized into a type abstraction hierarchy. This hierarchy allows a word such as "mass" to be relaxed to the more general abstract class "physobj.abnormal.findin- g.lesion". Within this class, the terms "nodule", and "cyst" are included. In this way, the restrictions of what types of words can fill specific roles of a given logical relation are relaxed. Relaxation can occur as many times as necessary in order to commit each word within the sentence to at least one possible logical relation role. Because there are hundreds of thousands of words in the English language, therefore it is likely that some words will not have appeared together in sentences in the training corpora and, as a result, there will be no template slot assignments for them. On the other hand, because the number of unique semantic classes is on the order of 500 and the number of syntactic classes is on the order of 15, these numbers are both finite sufficiently small to ensure that at least a default role can be supplied for any unforeseen combination of words. Moreover, of the 500 semantic classes, many of these are parts of hierarchies. For example, "mass" is classified as "physobj.abnorm.finding.lesion," a class which can be broadened to "physobj.abnorm.finding" or broadened even further to the class of all "physobj.abnorm." Accordingly, the builders of the database can choose how they might want to balance accuracy and detail in building this concept database (Taira [0111-0112] & Fig. 7 iterations, Fig. 12 relaxing and broadening terms).

Therefore, it would have been obvious to one of ordinary skill in the art at the time of the invention to modify the system of Zhou to incorporate moving iteratively up the semantic hierarchy of the constraint if the pattern entry after constraint relaxation is determined not to have a valid form comprises moving up the semantic hierarchy of the constraint, relaxing the constraint further, and returning to determining if the pattern entry after constraint relaxation has a valid form as taught by Taira to allow for identification of the best match of input NER data with various interpretations, wherein relaxation can occur as many times as necessary in order to commit each word within the sentence to at least one possible logical relation role (Taira [0111-0112] & Fig. 7 iterations, Fig. 12 relaxing and broadening terms)., thereby valid constraint forms as taught by Zhou can be improved to include a broad and narrow use of NER entries by continuously broadening (if necessary) through disambiguation, frequency of occurrence, and constraint relaxation (Taira Fig. 7 for example iterative elements 730, 740, and 750, having relaxation applied [0111-0112]).

Re claims 3 and 14, Zhou teaches valid and invalid constraint forms in a hierarchy of chunks (page 475 col. 2).

However, Zhou fails to teach determining if a constraint in the pattern entry, after relaxation, also has a valid form and moving iteratively up the semantic hierarchy of the constraint if the constraint in the pattern entry after constraint relaxation is determined not to have a valid form

Taira initially teaches matching text segments where statistics are accumulated on the frequency of occurrence within the training set of how often a word class is associated with a given logical relation role as head, relation, or value. This model is created to allow for the rapid indexing of the possible logical relations in which each word might be involved. Thus, information is gathered about the probabilistic context-free role of the particle in constructing logical relations. This initial characterization is used as a first pass filter to determine possible candidate logical relations for a given word (Taira [0107]).

Further, Taira teaches a degree of deformability is desirable in a template matching algorithm whose goal is high recall. For example, some semantic smoothing can be performed so that individual slot value constraints of the logical relation template are relaxed (Taira [0108]).

Furthermore, Taira teaches the relaxing of constraints to achieve a match if frequency of occurrence does not produce an exact match, wherein Taira teaches [0111] One problem inherent in creating a template matching algorithm automatically from a finite set of training data is that it is unlikely that every combination of words that might be used in a natural language medical report will be addressed. Unforeseen combinations of words will leave gaps in the systems ability to identify all possible logical relation word pairs consisting of heads and values or triplets consisting of heads, relations, and values. To account for this problem, if there is no logical relation slot assignment for a particular words, one will be estimated using a graded relaxation of the syntactic and semantic properties of the word. FIG. 12 illustrates the mechanism used

for relaxation. The semantic classes assigned to words is organized into a type abstraction hierarchy. This hierarchy allows a word such as "mass" to be relaxed to the more general abstract class "physobj.abnormal.findin- g.lesion". Within this class, the terms "nodule", and "cyst" are included. In this way, the restrictions of what types of words can fill specific roles of a given logical relation are relaxed. Relaxation can occur as many times as necessary in order to commit each word within the sentence to at least one possible logical relation role. Because there are hundreds of thousands of words in the English language, therefore it is likely that some words will not have appeared together in sentences in the training corpora and, as a result, there will be no template slot assignments for them. On the other hand, because the number of unique semantic classes is on the order of 500 and the number of syntactic classes is on the order of 15, these numbers are both finite sufficiently small to ensure that at least a default role can be supplied for any unforeseen combination of words. Moreover, of the 500 semantic classes, many of these are parts of hierarchies. For example, "mass" is classified as "physobj.abnorm.finding.lesion," a class which can be broadened to "physobj.abnorm.finding" or broadened even further to the class of all "physobj.abnorm." Accordingly, the builders of the database can choose how they might want to balance accuracy and detail in building this concept database (Taira [0111-0112] & Fig. 7 iterations, Fig. 12 relaxing and broadening terms).

Therefore, it would have been obvious to one of ordinary skill in the art at the time of the invention to modify the system of Zhou to incorporate determining if a constraint in the pattern entry, after relaxation, also has a valid form and moving

iteratively up the semantic hierarchy of the constraint if the constraint in the pattern entry after constraint relaxation is determined not to have a valid form as taught by Taira to allow for identification of the best match of input NER data with various interpretations, wherein relaxation can occur as many times as necessary in order to commit each word within the sentence to at least one possible logical relation role (Taira [0111-0112] & Fig. 7 iterations, Fig. 12 relaxing and broadening terms)., thereby valid constraint forms as taught by Zhou can be improved to include a broad and narrow use of NER entries by continuously broadening (if necessary) through disambiguation, frequency of occurrence, and constraint relaxation (Taira Fig. 7 for example iterative elements 730, 740, and 750, having relaxation applied [0111-0112]).

Re claims 4 and 15, Zhou teaches valid and invalid constraint forms in a hierarchy of chunks (page 475 col. 2).

However, Zhou fails to teach moving iteratively up the semantic hierarchy of the constraint if the constraint in the pattern entry after constraint relaxation is determined not to have a valid form comprises moving up the semantic hierarchy of the constraint, relaxing the constraint further, and returning to determining if a constraint in the pattern entry after constraint relaxation has a valid form.

Taira initially teaches matching text segments where statistics are accumulated on the frequency of occurrence within the training set of how often a word class is associated with a given logical relation role as head, relation, or value. This model is created to allow for the rapid indexing of the possible logical relations in which each

word might be involved. Thus, information is gathered about the probabilistic context-free role of the particle in constructing logical relations. This initial characterization is used as a first pass filter to determine possible candidate logical relations for a given word (Taira [0107]).

Further, Taira teaches a degree of deformability is desirable in a template matching algorithm whose goal is high recall. For example, some semantic smoothing can be performed so that individual slot value constraints of the logical relation template are relaxed (Taira [0108]).

Furthermore, Taira teaches the relaxing of constraints to achieve a match if frequency of occurrence does not produce an exact match, wherein Taira teaches [0111] One problem inherent in creating a template matching algorithm automatically from a finite set of training data is that it is unlikely that every combination of words that might be used in a natural language medical report will be addressed. Unforeseen combinations of words will leave gaps in the systems ability to identify all possible logical relation word pairs consisting of heads and values or triplets consisting of heads, relations, and values. To account for this problem, if there is no logical relation slot assignment for a particular words, one will be estimated using a graded relaxation of the syntactic and semantic properties of the word. FIG. 12 illustrates the mechanism used for relaxation. The semantic classes assigned to words is organized into a type abstraction hierarchy. This hierarchy allows a word such as "mass" to be relaxed to the more general abstract class "physobj.abnormal.findin- g.lesion". Within this class, the terms "nodule", and "cyst" are included. In this way, the restrictions of what types of

words can fill specific roles of a given logical relation are relaxed. Relaxation can occur as many times as necessary in order to commit each word within the sentence to at least one possible logical relation role. Because there are hundreds of thousands of words in the English language, therefore it is likely that some words will not have appeared together in sentences in the training corpora and, as a result, there will be no template slot assignments for them. On the other hand, because the number of unique semantic classes is on the order of 500 and the number of syntactic classes is on the order of 15, these numbers are both finite sufficiently small to ensure that at least a default role can be supplied for any unforeseen combination of words. Moreover, of the 500 semantic classes, many of these are parts of hierarchies. For example, "mass" is classified as "physobj.abnorm.finding.lesion," a class which can be broadened to "physobj.abnorm.finding" or broadened even further to the class of all "physobj.abnorm." Accordingly, the builders of the database can choose how they might want to balance accuracy and detail in building this concept database (Taira [0111-0112] & Fig. 7 iterations, Fig. 12 relaxing and broadening terms).

Therefore, it would have been obvious to one of ordinary skill in the art at the time of the invention to modify the system of Zhou to incorporate moving iteratively up the semantic hierarchy of the constraint if the constraint in the pattern entry after constraint relaxation is determined not to have a valid form comprises moving up the semantic hierarchy of the constraint, relaxing the constraint further, and returning to determining if a constraint in the pattern entry after constraint relaxation has a valid form as taught by Taira to allow for identification of the best match of input NER data with

various interpretations, wherein relaxation can occur as many times as necessary in order to commit each word within the sentence to at least one possible logical relation role (Taira [0111-0112] & Fig. 7 iterations, Fig. 12 relaxing and broadening terms)., thereby valid constraint forms as taught by Zhou can be improved to include a broad and narrow use of NER entries by continuously broadening (if necessary) through disambiguation, frequency of occurrence, and constraint relaxation (Taira Fig. 7 for example iterative elements 730, 740, and 750, having relaxation applied [0111-0112]).

Re claim 5, Zhou fails to teach a method according to claim 1, wherein if a constraint is relaxed, the constraint is dropped entirely from the pattern entry if the relaxation reaches the root of the semantic hierarchy.

Taira teaches a degree of deformability is desirable in a template matching algorithm whose goal is high recall. For example, some semantic smoothing can be performed so that individual slot value constraints of the logical relation template are relaxed (Taira [0108]).

Further, Taira teaches the relaxing of constraints to achieve a match if frequency of occurrence does not produce an exact match, wherein Taira teaches [0111] One problem inherent in creating a template matching algorithm automatically from a finite set of training data is that it is unlikely that every combination of words that might be used in a natural language medical report will be addressed. Unforeseen combinations of words will leave gaps in the systems ability to identify all possible logical relation word pairs consisting of heads and values or triplets consisting of heads, relations, and

values. To account for this problem, if there is no logical relation slot assignment for a particular words, one will be estimated using a graded relaxation of the syntactic and semantic properties of the word. FIG. 12 illustrates the mechanism used for relaxation. The semantic classes assigned to words is organized into a type abstraction hierarchy. This hierarchy allows a word such as "mass" to be relaxed to the more general abstract class "physobj.abnormal.findin- g.lesion". Within this class, the terms "nodule", and "cyst" are included. In this way, the restrictions of what types of words can fill specific roles of a given logical relation are relaxed. Relaxation can occur as many times as necessary in order to commit each word within the sentence to at least one possible logical relation role. Because there are hundreds of thousands of words in the English language, therefore it is likely that some words will not have appeared together in sentences in the training corpora and, as a result, there will be no template slot assignments for them. On the other hand, because the number of unique semantic classes is on the order of 500 and the number of syntactic classes is on the order of 15, these numbers are both finite sufficiently small to ensure that at least a default role can be supplied for any unforeseen combination of words. Moreover, of the 500 semantic classes, many of these are parts of hierarchies. For example, "mass" is classified as "physobj.abnorm.finding.lesion," a class which can be broadened to "physobj.abnorm.finding" or broadened even further to the class of all "physobj.abnorm." Accordingly, the builders of the database can choose how they might want to balance accuracy and detail in building this concept database (Taira [0111-0112] & Fig. 7 iterations, Fig. 12 relaxing and broadening terms).

Therefore, it would have been obvious to one of ordinary skill in the art at the time of the invention to modify the system of Zhou to incorporate if a constraint is relaxed, the constraint is dropped entirely from the pattern entry if the relaxation reaches the root of the semantic hierarchy as taught by Taira to allow for identification of the best match of input NER data with various interpretations, wherein relaxation can occur as many times as necessary in order to commit each word within the sentence to at least one possible logical relation role (Taira [0111-0112] & Fig. 7 iterations, Fig. 12 relaxing and broadening terms)., thereby valid constraint forms as taught by Zhou can be improved to include a broad and narrow use of NER entries by continuously broadening (if necessary) through disambiguation, frequency of occurrence, and constraint relaxation (Taira Fig. 7 for example iterative elements 730, 740, and 750, having relaxation applied [0111-0112]).

Re claim 6, Zhou fails to teach a method according to claim 1, further comprising terminating if a near optimal frequently occurring pattern entry is reached to replace the initial pattern entry.

Taira initially teaches matching text segments where statistics are accumulated on the frequency of occurrence within the training set of how often a word class is associated with a given logical relation role as head, relation, or value. This model is created to allow for the rapid indexing of the possible logical relations in which each word might be involved. Thus, information is gathered about the probabilistic context-free role of the particle in constructing logical relations. This initial characterization is

used as a first pass filter to determine possible candidate logical relations for a given word (Taira [0107]).

Further, Taira teaches a degree of deformability is desirable in a template matching algorithm whose goal is high recall. For example, some semantic smoothing can be performed so that individual slot value constraints of the logical relation template are relaxed (Taira [0108]).

Furthermore, Taira teaches the relaxing of constraints to achieve a match if frequency of occurrence does not produce an exact match, wherein Taira teaches [0111] One problem inherent in creating a template matching algorithm automatically from a finite set of training data is that it is unlikely that every combination of words that might be used in a natural language medical report will be addressed. Unforeseen combinations of words will leave gaps in the systems ability to identify all possible logical relation word pairs consisting of heads and values or triplets consisting of heads, relations, and values. To account for this problem, if there is no logical relation slot assignment for a particular words, one will be estimated using a graded relaxation of the syntactic and semantic properties of the word. FIG. 12 illustrates the mechanism used for relaxation. The semantic classes assigned to words is organized into a type abstraction hierarchy. This hierarchy allows a word such as "mass" to be relaxed to the more general abstract class "physobj.abnormal.findin- g.lesion". Within this class, the terms "nodule", and "cyst" are included. In this way, the restrictions of what types of words can fill specific roles of a given logical relation are relaxed. Relaxation can occur as many times as necessary in order to commit each word within the sentence to at

least one possible logical relation role. Because there are hundreds of thousands of words in the English language, therefore it is likely that some words will not have appeared together in sentences in the training corpora and, as a result, there will be no template slot assignments for them. On the other hand, because the number of unique semantic classes is on the order of 500 and the number of syntactic classes is on the order of 15, these numbers are both finite sufficiently small to ensure that at least a default role can be supplied for any unforeseen combination of words. Moreover, of the 500 semantic classes, many of these are parts of hierarchies. For example, "mass" is classified as "physobj.abnorm.finding.lesion," a class which can be broadened to "physobj.abnorm.finding" or broadened even further to the class of all "physobj.abnorm." Accordingly, the builders of the database can choose how they might want to balance accuracy and detail in building this concept database (Taira [0111-0112] & Fig. 7 iterations, Fig. 12 relaxing and broadening terms).

Therefore, it would have been obvious to one of ordinary skill in the art at the time of the invention to modify the system of Zhou to incorporate terminating if a near optimal frequently occurring pattern entry is reached to replace the initial pattern entry as taught by Taira to allow for identification of the best match of input NER data with various interpretations, wherein relaxation can occur as many times as necessary in order to commit each word within the sentence to at least one possible logical relation role (Taira [0111-0112] & Fig. 7 iterations, Fig. 12 relaxing and broadening terms)., thereby valid constraint forms as taught by Zhou can be improved to include a broad and narrow use of NER entries by continuously broadening (if necessary) through

disambiguation, frequency of occurrence, and constraint relaxation (Taira Fig. 7 for example iterative elements 730, 740, and 750, having relaxation applied [0111-0112]).

Re claim 7, Zhou teaches a method according to claim 1, further comprising selecting the initial pattern entry for back-off modeling if it is not a frequently occurring pattern entry in a lexicon.

However, Zhou fails to teach if it is not a frequently occurring pattern entry in a lexicon

Taira initially teaches matching text segments where statistics are accumulated on the frequency of occurrence within the training set of how often a word class is associated with a given logical relation role as head, relation, or value. This model is created to allow for the rapid indexing of the possible logical relations in which each word might be involved. Thus, information is gathered about the probabilistic context-free role of the particle in constructing logical relations. This initial characterization is used as a first pass filter to determine possible candidate logical relations for a given word (Taira [0107]).

Further, Taira teaches a degree of deformability is desirable in a template matching algorithm whose goal is high recall. For example, some semantic smoothing can be performed so that individual slot value constraints of the logical relation template are relaxed (Taira [0108]).

Furthermore, Taira teaches the relaxing of constraints to achieve a match if frequency of occurrence does not produce an exact match, as recited in claim 1 (Taira [0111-0112] & Fig. 7 iterations, Fig. 12 relaxing and broadening terms).

Therefore, it would have been obvious to one of ordinary skill in the art at the time of the invention to modify the system of Zhou to incorporate selecting the initial pattern entry for back-off modeling if it is not a frequently occurring pattern entry in a lexicon if it is not a frequently occurring pattern entry in a lexicon as taught by Taira to allow for identification of the best match of input NER data with various interpretations, wherein relaxation can occur as many times as necessary in order to commit each word within the sentence to at least one possible logical relation role (Taira [0111-0112] & Fig. 7 iterations, Fig. 12 relaxing and broadening terms)., thereby valid constraint forms as taught by Zhou can be improved to include a broad and narrow use of NER entries by continuously broadening (if necessary) through disambiguation, frequency of occurrence, and constraint relaxation (Taira Fig. 7 for example iterative elements 730, 740, and 750, having relaxation applied [0111-0112]).

Re claim 8, Zhou teaches a method of inducing patterns in a pattern lexicon (page 475 section 3 and page 476 section 3.1) comprising a

However, Zhou fails to teach a plurality of initial pattern entries with associated occurrence frequencies, the method comprising:

identifying one or more initial pattern entries in the lexicon with lower occurrence frequencies;

relaxing one or more constraints of individual ones of the identified one or more initial pattern entries to broaden the coverage of the identified one or more initial pattern entries.

Taira initially teaches matching text segments where statistics are accumulated on the frequency of occurrence within the training set of how often a word class is associated with a given logical relation role as head, relation, or value. This model is created to allow for the rapid indexing of the possible logical relations in which each word might be involved. Thus, information is gathered about the probabilistic context-free role of the particle in constructing logical relations. This initial characterization is used as a first pass filter to determine possible candidate logical relations for a given word (Taira [0107]).

Further, Taira teaches a degree of deformability is desirable in a template matching algorithm whose goal is high recall. For example, some semantic smoothing can be performed so that individual slot value constraints of the logical relation template are relaxed (Taira [0108]).

Furthermore, Taira teaches the relaxing of constraints to achieve a match if frequency of occurrence does not produce an exact match, as recited in claim 1 (Taira [0111-0112] & Fig. 7 iterations, Fig. 12 relaxing and broadening terms).

Therefore, it would have been obvious to one of ordinary skill in the art at the time of the invention to modify the system of Zhou to incorporate a plurality of initial pattern entries with associated occurrence frequencies, the method comprising identifying one or more initial pattern entries in the lexicon with lower occurrence

frequencies and relaxing one or more constraints of individual ones of the identified one or more initial pattern entries to broaden the coverage of the identified one or more initial pattern entriesas taught by Taira to allow for identification of the best match of input NER data with various interpretations, wherein relaxation can occur as many times as necessary in order to commit each word within the sentence to at least one possible logical relation role (Taira [0111-0112] & Fig. 7 iterations, Fig. 12 relaxing and broadening terms)., thereby valid constraint forms as taught by Zhou can be improved to include a broad and narrow use of NER entries by continuously broadening (if necessary) through disambiguation, frequency of occurrence, and constraint relaxation (Taira Fig. 7 for example iterative elements 730, 740, and 750, having relaxation applied [0111-0112]).

Re claim 9, Zhou teaches a method according to claim 8, further comprising creating the pattern lexicon of initial pattern entries from a training corpus (page 475 section 3 and page 476 section 3.1).

Re claim 10, Zhou fails to teach a method according to claim 8, further comprising merging individual ones of the constraint relaxed initial pattern entries with similar pattern entries in the lexicon to form a more compact pattern lexicon.

Taira teaches that because a separate maximum entropy classifier is developed for each type of logical relation, each classifier is more compact and, overall, the system is more accurate. For example, the classifier for extracting the logical relation corresponding to the predicate "hasArticle" is designed separately from that of say the predicate "hasSize." This allows only features significant for discriminating the existence of these logical relations to be encapsulated within the specific classifier. Additionally, it allows one to easily collect training examples for any defined logical relation using a retrieval engine. Thus, even though the prevalence of "hasArticle" logical relations is much greater than say "hasSize", one can easily get enough training examples for "hasSize" by simply retrieving over a larger corpus of reports. Thus, even for fairly infrequent types of logical relations, such as "hasShape," one can assemble a large number of training data. Moreover, the training data for a given logical relation requires only tagging logical relation instances of the given type within a training sentence and no others, saving programmer time (Taira [0133]).

Taira initially teaches matching text segments where statistics are accumulated on the frequency of occurrence within the training set of how often a word class is associated with a given logical relation role as head, relation, or value. This model is created to allow for the rapid indexing of the possible logical relations in which each word might be involved. Thus, information is gathered about the probabilistic context-free role of the particle in constructing logical relations. This initial characterization is used as a first pass filter to determine possible candidate logical relations for a given word (Taira [0107]).

Further, Taira teaches a degree of deformability is desirable in a template matching algorithm whose goal is high recall. For example, some semantic smoothing can be performed so that individual slot value constraints of the logical relation template are relaxed (Taira [0108]).

Furthermore, Taira teaches the relaxing of constraints to achieve a match if frequency of occurrence does not produce an exact match, as recited in claim 1 (Taira [0111-0112] & Fig. 7 iterations, Fig. 12 relaxing and broadening terms).

Therefore, it would have been obvious to one of ordinary skill in the art at the time of the invention to modify the system of Zhou to incorporate merging individual ones of the constraint relaxed initial pattern entries with similar pattern entries in the lexicon to form a more compact pattern lexicon as taught by Taira to allow for identification of the best match of input NER data with various interpretations, wherein relaxation can occur as many times as necessary in order to commit each word within the sentence to at least one possible logical relation role (Taira [0111-0112] & Fig. 7 iterations, Fig. 12 relaxing and broadening terms)., thereby valid constraint forms as taught by Zhou can be improved to include a broad and narrow use of NER entries by continuously broadening (if necessary) through disambiguation, frequency of occurrence, and constraint relaxation (Taira Fig. 7 for example iterative elements 730, 740, and 750, having relaxation applied [0111-0112]), wherein a compact classification scheme can emerge that requires only tagging logical relation instances of the given type within a training sentence and no others, saving programmer time through broadened patterns (Taira [0133]).

Re claim 11, method according to claim 9, wherein the entries in the compact pattern lexicon are generalized as much as possible within a given similarity threshold

Taira initially teaches matching text segments where statistics are accumulated on the frequency of occurrence within the training set of how often a word class is associated with a given logical relation role as head, relation, or value. This model is created to allow for the rapid indexing of the possible logical relations in which each word might be involved. Thus, information is gathered about the probabilistic context-free role of the particle in constructing logical relations. This initial characterization is used as a first pass filter to determine possible candidate logical relations for a given word (Taira [0107]).

Further, Taira teaches a degree of deformability is desirable in a template matching algorithm whose goal is high recall. For example, some semantic smoothing can be performed so that individual slot value constraints of the logical relation template are relaxed (Taira [0108]).

Furthermore, Taira teaches the relaxing of constraints to achieve a match if frequency of occurrence does not produce an exact match, as recited in claim 1 (Taira [0111-0112] & Fig. 7 iterations, Fig. 12 relaxing and broadening terms).

Therefore, it would have been obvious to one of ordinary skill in the art at the time of the invention to modify the system of Zhou to incorporate entries in the compact pattern lexicon are generalized as much as possible within a given similarity threshold as taught by Taira to allow for identification of the best match of input NER data with

various interpretations, wherein relaxation can occur as many times as necessary in order to commit each word within the sentence to at least one possible logical relation role (Taira [0111-0112] & Fig. 7 iterations, Fig. 12 relaxing and broadening terms)., thereby valid constraint forms as taught by Zhou can be improved to include a broad and narrow use of NER entries by continuously broadening (if necessary) through disambiguation, frequency of occurrence, and constraint relaxation (Taira Fig. 7 for example iterative elements 730, 740, and 750, having relaxation applied [0111-0112]).

Re claim 16, Zhou teaches a decoding process in a rich feature space comprising a method according to claim 1 (page 479 col. 2).

Re claim 17, Zhou teaches a training process in a rich feature space comprising a method according to claim 8 (page 479 col. 2).

Re claim 18, Zhou teaches a system for recognizing and classifying named entities within a text, comprising:

feature extraction means for extracting various features from a document; (page 479 col. 2)

recognition kernel means to recognize and classify named entities using a Hidden Markov Model (page 479 col. 2);

back-off modeling means for back-off modeling by constraint relaxation to deal with data sparseness in a rich feature space (page 478 col. 1).

However, Zhou fails to teach constraint relaxation

Taira initially teaches matching text segments where statistics are accumulated on the frequency of occurrence within the training set of how often a word class is associated with a given logical relation role as head, relation, or value. This model is created to allow for the rapid indexing of the possible logical relations in which each word might be involved. Thus, information is gathered about the probabilistic context-free role of the particle in constructing logical relations. This initial characterization is used as a first pass filter to determine possible candidate logical relations for a given word (Taira [0107]).

Further, Taira teaches a degree of deformability is desirable in a template matching algorithm whose goal is high recall. For example, some semantic smoothing can be performed so that individual slot value constraints of the logical relation template are relaxed (Taira [0108]).

Furthermore, Taira teaches the relaxing of constraints to achieve a match if frequency of occurrence does not produce an exact match, as recited in claim 1 (Taira [0111-0112] & Fig. 7 iterations, Fig. 12 relaxing and broadening terms).

Therefore, it would have been obvious to one of ordinary skill in the art at the time of the invention to modify the system of Zhou to incorporate back-off modeling means for back-off modeling by constraint relaxation as taught by Taira to allow for identification of the best match of input NER data with various interpretations, wherein relaxation can occur as many times as necessary in order to commit each word within the sentence to at least one possible logical relation role (Taira [0111-0112] & Fig. 7

iterations, Fig. 12 relaxing and broadening terms)., thereby valid constraint forms as taught by Zhou can be improved to include a broad and narrow use of NER entries by continuously broadening (if necessary) through disambiguation, frequency of occurrence, and constraint relaxation (Taira Fig. 7 for example iterative elements 730, 740, and 750, having relaxation applied [0111-0112]), wherein a compact classification scheme can emerge that requires only tagging logical relation instances of the given type within a training sentence and no others, saving programmer time through broadened patterns (Taira [0133]).

Re claim 19, Zhou teaches a system according to claim 18, wherein the back-off modeling means is operable to provide a method of back-off modeling according to claim 1 (page 478 col. 1).

Re claim 20, Zhou fails to teach a system according to claim 18, further comprising a pattern induction means for inducing frequently occurring patterns

Taira initially teaches matching text segments where statistics are accumulated on the frequency of occurrence within the training set of how often a word class is associated with a given logical relation role as head, relation, or value. This model is created to allow for the rapid indexing of the possible logical relations in which each word might be involved. Thus, information is gathered about the probabilistic context-free role of the particle in constructing logical relations. This initial characterization is

used as a first pass filter to determine possible candidate logical relations for a given word (Taira [0107]).

Further, Taira teaches a degree of deformability is desirable in a template matching algorithm whose goal is high recall. For example, some semantic smoothing can be performed so that individual slot value constraints of the logical relation template are relaxed (Taira [0108]).

Furthermore, Taira teaches the relaxing of constraints to achieve a match if frequency of occurrence does not produce an exact match, as recited in claim 1 (Taira [0111-0112] & Fig. 7 iterations, Fig. 12 relaxing and broadening terms).

Therefore, it would have been obvious to one of ordinary skill in the art at the time of the invention to modify the system of Zhou to incorporate a pattern induction means for inducing frequently occurring patterns as taught by Taira to allow for identification of the best match of input NER data with various interpretations, wherein relaxation can occur as many times as necessary in order to commit each word within the sentence to at least one possible logical relation role (Taira [0111-0112] & Fig. 7 iterations, Fig. 12 relaxing and broadening terms)., thereby valid constraint forms as taught by Zhou can be improved to include a broad and narrow use of NER entries by continuously broadening (if necessary) through disambiguation, frequency of occurrence, and constraint relaxation (Taira Fig. 7 for example iterative elements 730, 740, and 750, having relaxation applied [0111-0112]), wherein a compact classification scheme can emerge that requires only tagging logical relation instances of the given

type within a training sentence and no others, saving programmer time through broadened patterns (Taira [0133]).

Re claim 21, Zhou teaches a system according to claim 20, wherein the pattern induction means is operable to provide a method of inducing patterns according to claim 8 (page 478 col. 1).

Re claim 22, Zhou teaches a system according to claim 18, wherein said various features are extracted from words within the text and the discourse of the text, and comprise one or more of:

a) deterministic features of words, including capitalization or digitalization (abstract)

b) semantic features of trigger words (abstract);

c) gazetteer features, which determine whether and how the current word string appears in a gazetteer list; (page 476 section 3.1)

d) discourse features, which deal with the phenomena of name alias (page 476 section 3.2);

e) the words themselves (page 478 section 4.1).

Re claim 23, Zhou teaches a feature set for use in back-off modeling in a Hidden Markov Model, during named entity recognition (page 479 col. 2), wherein the feature sets are arranged hierarchically (page 476 section 3.2) to allow for data sparseness

However, Zhou fails to teach feature sets are arranged hierarchically to allow for data sparseness

Taira initially teaches matching text segments where statistics are accumulated on the frequency of occurrence within the training set of how often a word class is associated with a given logical relation role as head, relation, or value. This model is created to allow for the rapid indexing of the possible logical relations in which each word might be involved. Thus, information is gathered about the probabilistic context-free role of the particle in constructing logical relations. This initial characterization is used as a first pass filter to determine possible candidate logical relations for a given word (Taira [0107]).

Further, Taira teaches a degree of deformability is desirable in a template matching algorithm whose goal is high recall. For example, some semantic smoothing can be performed so that individual slot value constraints of the logical relation template are relaxed (Taira [0108]).

Furthermore, Taira teaches the relaxing of constraints to achieve a match if frequency of occurrence does not produce an exact match, as recited in claim 1 (Taira [0111-0112] & Fig. 7 iterations, Fig. 12 relaxing and broadening terms).

Therefore, it would have been obvious to one of ordinary skill in the art at the time of the invention to modify the system of Zhou to incorporate feature sets that are

arranged hierarchically to allow for data sparseness as taught by Taira to allow for identification of the best match of input NER data with various interpretations, wherein relaxation can occur as many times as necessary in order to commit each word within the sentence to at least one possible logical relation role (Taira [0111-0112] & Fig. 7 iterations, Fig. 12 relaxing and broadening terms)., thereby valid constraint forms as taught by Zhou can be improved to include a broad and narrow use of NER entries by continuously broadening (if necessary) through disambiguation, frequency of occurrence, and constraint relaxation (Taira Fig. 7 for example iterative elements 730, 740, and 750, having relaxation applied [0111-0112]), wherein a compact classification scheme can emerge that requires only tagging logical relation instances of the given type within a training sentence and no others, saving programmer time through broadened patterns (Taira [0133]).

Conclusion

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Michael C. Colucci whose telephone number is (571)-270-1847. The examiner can normally be reached on 9:30 am - 6:00 pm, Monday-Friday.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Richemond Dorvil can be reached on (571)-272-7602. The fax phone

number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

/Michael C Colucci/
Examiner, Art Unit 2626
Patent Examiner
AU 2626
(571)-270-1847
Examiner FAX: (571)-270-2847
Michael.Colucci@uspto.gov

/Richemond Dorvil/
Supervisory Patent Examiner, Art Unit 2626